

Employing Standards in Support of Dynamic Assembly of SEEDS Components

A Community Oriented Federation-SEEDS Prototype Study

Final Report to be found on ESIP Federation Web site
www.esipfed.org, late March '03

Prototype Bases

- Desired model for reducing overhead in present more manual operations
- Reduction in turn-around-time, end-user cost
- Purpose in SEEDS, Distributed Data Management Framework to:
 - Maximize availability and utility of ESE products
 - Leverage community expertise, ideas, capabilities
 - Improve overall effectiveness of ESE-funded systems and services

Vol. 1, Strategic Evolution of Earth Science Enterprise Data Systems (SEEDS), Formulation Team Draft, December 2002

Study Project Characteristics

- Tested a community oriented process relative to data handling standards (systems approach)
- Interchange among ESE components and data distribution to end-users as a Framework
- Emphasizes adoption existing data interchange standards
- Evolved generalizable end-user standard
- Reduction in manual interventions
- Cooperation across diverse groups a potential SEEDS model

The FIND (Federation Interactive Network for Discovery)

- Major credit for the FIND, Federation Interoperability Group (FIG)
- Study project built on FIND
- FIND currently includes GCMD and Mercury
- FIND's Web-based indexing and metadata extraction and navigational interface tools an aid to discovery
- For Community Engagement efforts, Web-based extraction tools aid in building a richer query interface
- User search criteria, including spatial extent, can be passed from FIND to an ESIP application

Expansion of FIND Role

- Adopt standards, practices already in use
- A Framework, for integrated end-to-end activity
- Data access, content, and service descriptions
 - **Open GIS Consortium (OGC) Web Mapping and Coverage Servers (WMS/WCS)**
 - **WWW Consortium (W3C) Web Service Description Language (WSDL); Universal Description, Discovery, and Integration (UDDI)**
 - **International Standards Organization (ISO), Image Objects Descriptor (MPEG7)**
- Standards adaptation enabling of a value chain
- Greater technologies coordination

Integrating WMS/WCS with FIND

- Earth Science: many data formats, models; opr. systems
- OGC's Web Mapping Server standards specify how to send maps over the Internet
- Adds WMS/WCS 4-D Viewer, FIND directly accessible
- WMS servers installed at 9 ESIP sites
- Data other than imagery, addressed through use of WCS, having a rich metadata model
- Mercury data access and discovery services enhanced
- Metrics: sites using WMS, data access time improvement, community request-delivery, turn-around-time

Using Web Services Description Language (WSDL)

- WSDL an extensible markup language (XML) for describing a service operating on information or procedural languages
- Allows for encapsulating Earth science data services: searching, retrieving, subsetting, overlaying
- Once a service (eg. subsetting) is described by WSDL, can easily be registered with public or private UDDI registry
- UDDI, like the yellow pages for seeking services
- A metric for WSDL, number of services available in WSDL standard form

Using MPEG7 Content Descriptions

- Applies to description of metadata and content as used in multimedia, tools available lowering learning cost to use
- Replaces many proprietary formats, with which automated interpretations are impossible
- Data creators able to insert richer scientific information along with raw data for wider use/application of a data set
- IBM has developed and released three MPEG7 annotation tools, available IBM Alphaworks Web site
- Like MPEG1 (Direct TV) and MPEG2 (DVD), MPEG7 is expected to have broad impact on content distribution
- Metric: number data sets described using MPEG7 and number accesses to these data sets

Summary: WSDL, MPEG7

- Enable data and information services on the Web in forms to be further analyzed and processed, without manual intervention
- MPEG7, descriptors for shape, texture, and multi-channel images and image sequences (eg. Series of satellite photos). Can describe fairly complex objects in an image
- Main benefit found with WSDL and MPEG7, they adhere to emerging service and content description standards beyond those of the Federation
- Many free SW tools for WSDL, UDDI, and MPEG7, thus lowering new application development cost

Framework Tryout

- Potential Central CA Coast agricultural problem, grapevine pest, the Glassy-winged Sharpshooter (GWSS)
- Paralleled invasion of fire ants in Southeastern US, an earlier study applying a chain of Federation and non-Federation components
- Community Engagement experience used to develop Framework of standards, extending core FIND standards with expanded query tools
- Would balance the “push” model for the NASA seeded supply chain with a “pull” model to allow end user to initiate data and service streams (Fig. 1, to follow)
- Accommodated core and community interface issues through creation of Epi-Exchange data and service descriptor language

Enabling Data Availability and Processing for Public Use

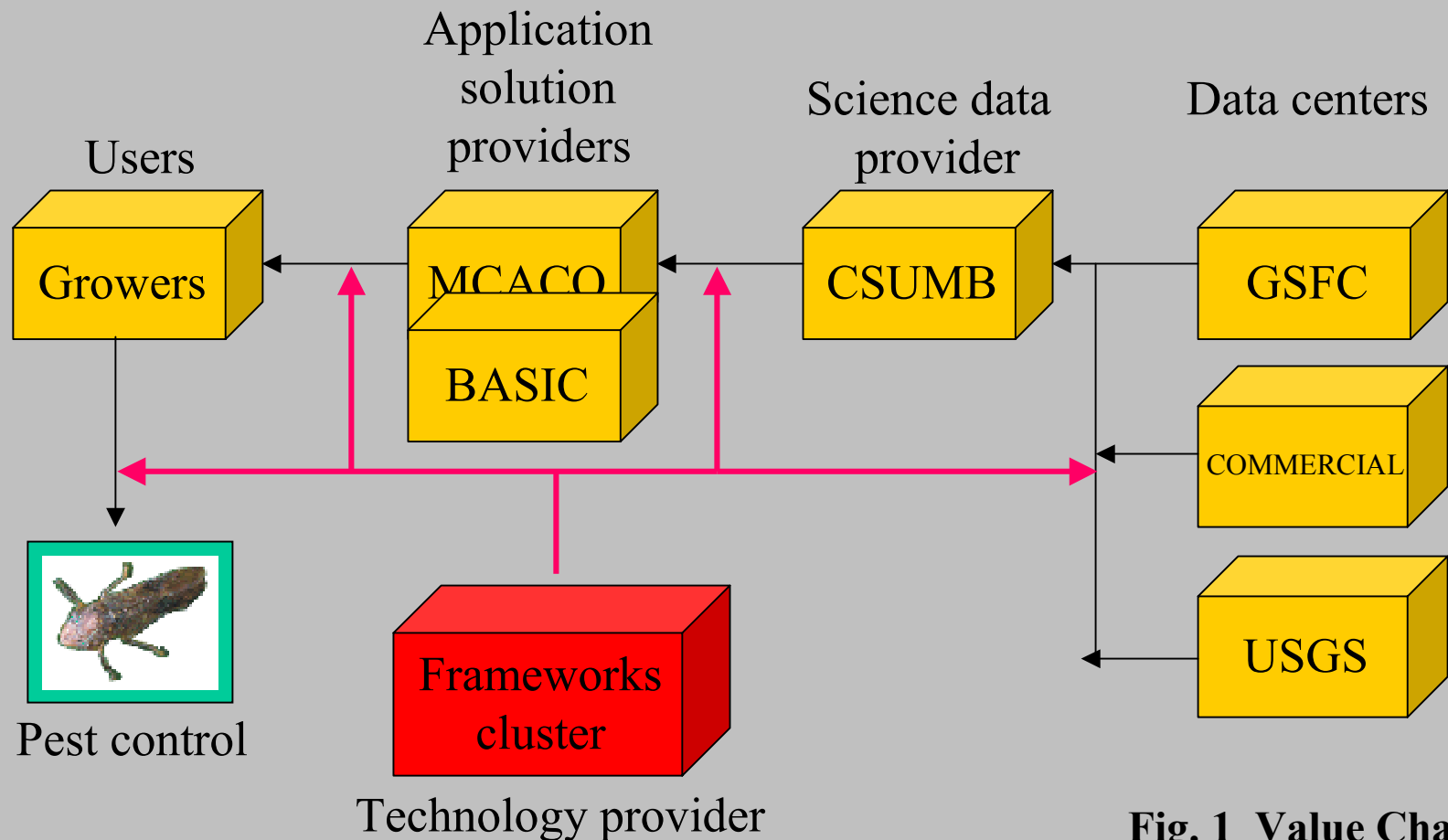


Fig. 1 Value Chain

Testing the Framework, with existing Community Data and Information

- **Community Need:** Management of the Glassy-winged Sharpshooter (GWSS), a vector spreading Pierce's Disease (PD)
- **Community:** BASIC (ESIP III) P.I., NASA Ames VINTAGE project, CSUMB SIVA Resources Center, Monterey Co. Ag. Comm. Office, County Vintners and Growers Assoc., Co. Farm Bureau, CA Central Coast PD Task Force
- **Desired Solution:** Digital Defense Map, perimeter determined by boundaries of NASA Ames digital color infrared (DCIR) image frame
- **RS Data Employed:** DCIR, section aerial photos, Landsat 7 ETM + scenes, USGS DEMs; Processes IP, GIS, GPS
- **Absent Data Sets:** High-res multispectral data, orthorectified imagery for accuracy, extended ground-truthing (per GPS)

Expanding the Framework in a Real-World Problem

- Epi-Exchange, an IBM product, represents the end user Framework interface
- Epi-Exchange the glue to create a data value chain
- All stakeholders along the data value chain may expect to benefit from Epi-Exchange
- Only a browser needed to access the product
- Problem-responsive elements developed by the Framework team (results Figs. 2&3 to follow, also see handout)
- Best metrics, the number of data value chains established, and, chain impacts (time saved, satisfaction ratings, etc.)

Fig. 2 An Example of RFQ Request on Epi-Exchange

The screenshot shows a Microsoft Internet Explorer window titled "RFQ Pest Invasion Model: Pilot Study Phase - epi-exchange". The address bar displays a URL from the Watson IBM QuickPlace. The page content is organized into a sidebar and a main body. The sidebar on the left contains a "Go" section with links like "Go Back", "RFQ Pest Invasion Model: Pilot Study Phase", "Instructions", "Room Index", "Negotiation", "Room Options", and "Room Security". Below this is a "Tools" section with links for "news: daily | weekly", "chat | notify | print | tutorial | help". The main body of the page features a header "RFQ Pest Invasion Model: Pilot Study Phase" with a timestamp "Vijay Sethia (vijay), 02/21/2003 - 03:03 PM". The content is divided into several sections: "Creator" (Bay Area Shared Information Consortium (BASIC), San Jose, CA; 408/345-1573), "Subject" (NAIC: entomology, aerial photography, satellite imagery, agriculture), "Objective" (Data sets sought for the Pest Invasion Model are natural color aerial photos, altitude about 12,000 feet, with an after scan GSD of 0.4 meters per pixel. Also desired is a scene of Landsat 7 ETM in the June-July timeframe (most recent available). Also desired is a USGS 7.5 minute (1:24,000) DEM, GSD of 30 meters. The final pilot study area dataset sought is that of a digital color-infrared (DCIR), 1-meter resolution, this to enable distinguishing among different crop types by row patterns and plant-bed width. Alternative datasets offering equal or greater usefulness may be proposed.), "Background" (To the data sets described in 4 above, local GIS specialists will add a vineyard ranch map overlay and also county land cover, roads and highways, city limits, and other feature overlays. Additionally, applying GIS tools, eight GWSS host/habitat type features will be identified on the base map. For reference purposes, bidders may wish to review the following Web sites pertaining to another pest invasion study involving the fire ant http://fireant.ifas.ufl.edu/), "Coverage" (Monterey County, CA is the pilot study local. Dataset coverage shall be the northern part of the mid-Salinas Valley to include the Valley floor and foothills from south of Chualar to just south of the Soledad State Correctional Facility. This is the northern most area where vineyards are planted in the Salinas Valley. US Highway 101 extends through the center of the coverage area. The main Quads involved are Rama, Gonzales, and Soledad. No area wide local measure is available at this time.), "Relation" (Standards applicable to this solicitation are WMS, WSDL, UDDI, and MPEG7. Pricing shall be on the basis of first approximation followed by best and final. Discounts and special offers may be indicated at either pricing point.), "Format" (The mode of delivery shall be electronic, either by Web or Disc. Delivery shall be not later than 30 days past the date of final negotiation, unless otherwise agreed.), "Pricing Range" (Per dataset cost should not exceed \$3,000 and the total cost for all spatial datasets here requested is not expected to exceed \$12,000.), and "Rights". The browser's status bar at the bottom shows "Done" and "Local intranet".

RFQ Pest Invasion Model: Pilot Study Phase
Vijay Sethia (vijay), 02/21/2003 - 03:03 PM

Creator
Bay Area Shared Information Consortium (BASIC), San Jose, CA; 408/345-1573

Subject
NAIC: entomology, aerial photography, satellite imagery, agriculture

Objective
Data sets sought for the Pest Invasion Model are natural color aerial photos, altitude about 12,000 feet, with an after scan GSD of 0.4 meters per pixel. Also desired is a scene of Landsat 7 ETM in the June-July timeframe (most recent available). Also desired is a USGS 7.5 minute (1:24,000) DEM, GSD of 30 meters. The final pilot study area dataset sought is that of a digital color-infrared (DCIR), 1-meter resolution, this to enable distinguishing among different crop types by row patterns and plant-bed width. Alternative datasets offering equal or greater usefulness may be proposed.

Background
To the data sets described in 4 above, local GIS specialists will add a vineyard ranch map overlay and also county land cover, roads and highways, city limits, and other feature overlays. Additionally, applying GIS tools, eight GWSS host/habitat type features will be identified on the base map. For reference purposes, bidders may wish to review the following Web sites pertaining to another pest invasion study involving the fire ant http://fireant.ifas.ufl.edu/

Coverage
Monterey County, CA is the pilot study local. Dataset coverage shall be the northern part of the mid-Salinas Valley to include the Valley floor and foothills from south of Chualar to just south of the Soledad State Correctional Facility. This is the northern most area where vineyards are planted in the Salinas Valley. US Highway 101 extends through the center of the coverage area. The main Quads involved are Rama, Gonzales, and Soledad. No area wide local measure is available at this time.

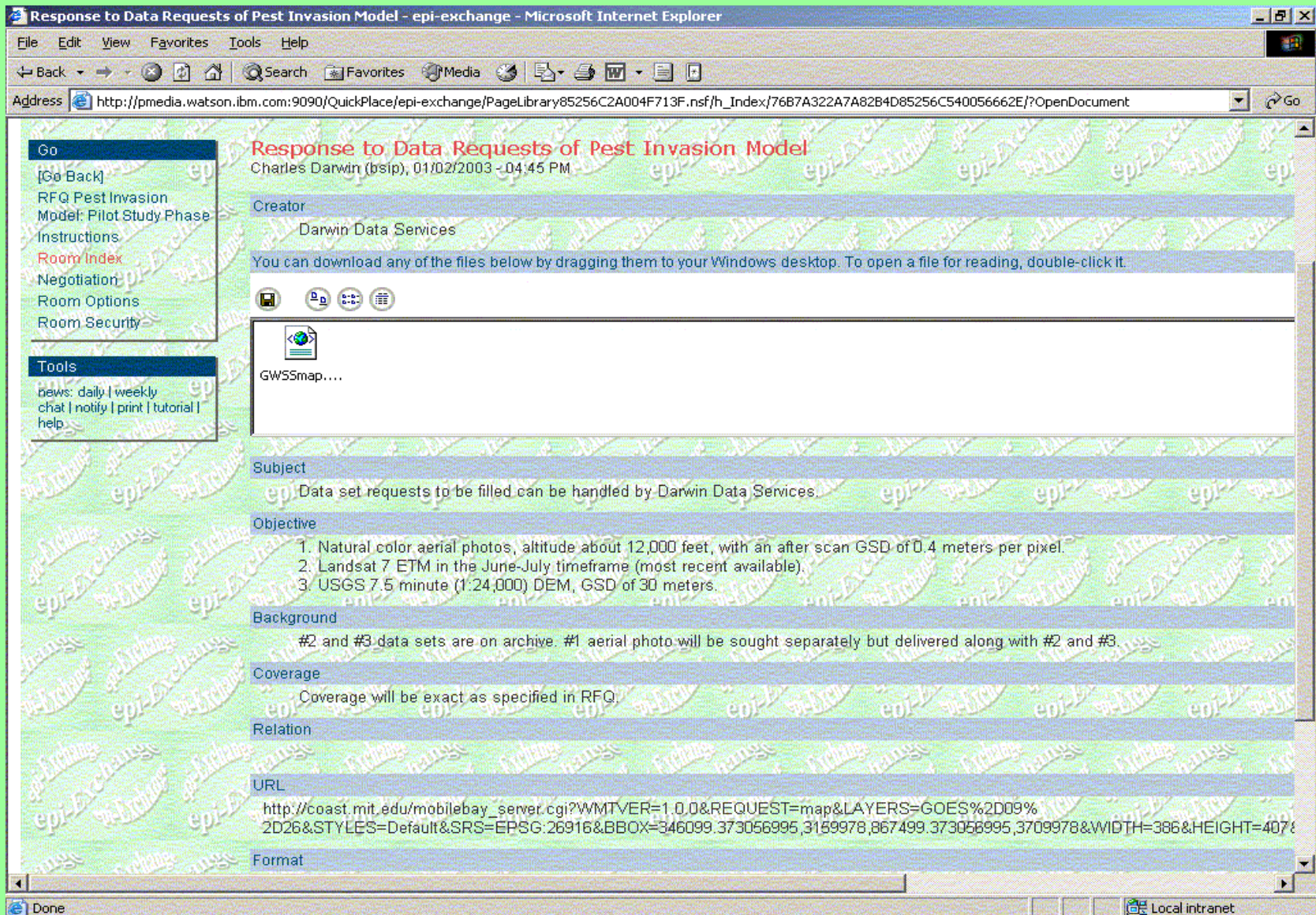
Relation
Standards applicable to this solicitation are WMS, WSDL, UDDI, and MPEG7. Pricing shall be on the basis of first approximation followed by best and final. Discounts and special offers may be indicated at either pricing point.

Format
The mode of delivery shall be electronic, either by Web or Disc. Delivery shall be not later than 30 days past the date of final negotiation, unless otherwise agreed.

Pricing Range
Per dataset cost should not exceed \$3,000 and the total cost for all spatial datasets here requested is not expected to exceed \$12,000.

Rights

Fig. 3 An Example of Response to the RFQ on Epi-Exchange



Standards Framework Study Project

Credits

- **Federation Interactive Network for Discovery (FIND);**
Oak Ridge National Laboratory (ORNL DAAC) Tim Rhyne
- **Web Mapping Server/Web Coverage Servers (WMS/**
WCS); *Jet Propulsion Laboratory (Ocean ESIP) Rob Raskin*
- **Web Service Description Language (WSDL), Universal**
Description, Discovery, and Integration (UDDI), and
image objects descriptor MPEG7; *Epi-SPIRE IBM/*
JHU Yuan-chi Chang,, Howard Burrows
- **Core and Community Standards Connectivity;** *Bay*
Area Shared Information Consortium (BASIC), P.I.,
with Epi-SPIRE IBM/JHU Dave Etter, Howard Burrows